Appendice 8

Protein-Protein Interaction in the Light of the Maximum Ordinality Principle

Corrado Giannantoni

ENEA's Researcher and Consultant of
Duchenne Parent Project Onlus
Rome, Italy
email: c.giannantoni@parentproject.it, corrado.giannantoni@tin.it

***Abstract* – The present paper is aimed at showing how it is possible to obtain Protein-Protein Interaction (PPI) in explicit formal terms, when this process is modeled by adopting the Maximum Ordinality Principle (MOP) as the basic reference criterion.**

Keywords-Protein-Protein Interaction (PPI); Maximum Ordinality Principle (MOP); Molecular Docking; Drug Design.

I. INTRODUCTION

Protein-Protein Interaction (PPI) decisively represents a fundamental process in Pharmacology. However, in spite of its recognized importance, PPI has not manifested all its related potentialities yet, mainly because of the intrinsic unsolvability, in explicit terms, of the famous "Three-body Problem", as demonstrated by H. Poincaré in 1889 [1].

This result represents a strong limitation, because it is also valid in *Protein Dynamics*. Not only (and especially) in Protein Folding, but also in PPI.

On the other hand, the research for a numerical solution often overcomes the computation capacities of the most powerful computers at present available (10 Petaflops).

Under such conditions, about 40 different approaches to PPI have been proposed in Literature. All of them, however, always introduce some (more or less) marked approximations. For example, the two interacting proteins are sometimes modeled as they were "rigid bodies". Consequently, apart from the solution to some particular (and specific) cases of PPI, such approximations do not always lead to satisfactory general solutions, when compared with experimental data.

This is because, in the absence of an explicit formal solution to PPI problem, all the various approaches adopted are also characterized by a correlative absence of an effective *predictive capacity*. In particular, when the latter is referred to the three-dimensional configuration of the final compound.

In the framework of such a "state of the art", the main aim of this paper is to bring out the possibility of obtaining Protein-Protein Interaction, in a *fast* and *reliable* way, as the *formal solution to an N-body interaction problem*, when the process is modeled on the basis of the Maximum Ordinality Principle (MOP).

This is because, after having obtained the solution to the "Three-body Problem" in terms of *fractional incipient derivatives* [2], previously introduced in [3][4], the extension to the case of N bodies was obtained in the contest of the mathematical formulation of the MOP [5].

This result immediately suggested its application to Protein Folding [6][7] and, now, to the case of PPI.

In this respect, Section 2 will preliminary present the input/out of the mathematical model adopted. Section 3 will illustrate the solution process through an ostensive example. The related informatics advances will be discussed in Section 4, while Section 5 will consider a possible extension of the same approach to other pharmacological fields. Section 6 (devoted to the conclusion) will reconsider all the previous aspects in the light the basic principles of self-organizing systems of *ordinal nature*.

II. INPUT/OUPUT OF THE MATHEMATICAL MODEL

The formal enunciation of the MOP, with specific reference to biological problems, was presented in [6][7]. Such a formulation is able to facilitate the solution to the PPI problem because, when the amino acids of a protein are modeled as they were "atoms" of a macromolecule, the 3D structure of the protein can be obtained without necessarily knowing its primary structure (that is, the specific linear sequence of its amino acids).

In fact, it is sufficient:

i) to know the *total number* of amino acids (*N*);

ii) to assign three parameters $(\Sigma_{12}, \Phi_{12}, \Theta_{12})$ that define, in polar coordinates, the reciprocal positions of two *arbitrary* amino acids, understood as being *one sole* "*isolated" entity*. This is also the reason why the latter is referred to its own internal reference system;

iii) to assign, in addition, six appropriate parameters $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \psi_1, \psi_2, \psi_3)$, that define the internal *Relation Space* (RS) of the protein analyzed.

More specifically: $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ characterize the spatial orientation of the protein (understood as a whole), with respect to its internal reference axes; whereas (ψ_1, ψ_2, ψ_3) define the periodicities (along the three basic axes) of the mathematical solutions which "emerge" from the MOP.

These solutions are precisely those that give the positions of all the amino acids with respect to the internal axes of the considered protein. In this way, the afore-mentioned solutions characterize any considered protein as a *unique*, *specific* and *irreducible* entity.

Under such conditions, each protein, precisely because modeled as a "self-organizing" system of *ordinal nature* (see Section 6), is also characterized by its own specific *self-organizing capacity*, whose *activity* can faithfully be represented by its associated "virtual work", defined (in polar coordinates) as

$$W = \sum_{j=2}^{N} \{(\rho_{1j}) + (\rho_{1j}\varphi_{1j}) + (\rho_{1j}\vartheta_{1j})\} \qquad (1)$$

where the subscript 1j indicates the couples of amino acids successively considered in the sum.

In the case of PPI, when there exists a given affinity between the interacting proteins, the resulting compound generally shows a "virtual work" W_3 that "exceeds" the sum of the "virtual works" W_1 and W_2 pertaining to the interacting proteins. Consequently, the *ratio* between such an excess of "virtual work"

$$\delta W = \{W_3 - (W_1 + W_2)\} \qquad (2)$$

and the sum of virtual works of the interacting proteins, that is

$$\delta W / (W_1 + W_2) \qquad (3)$$

can be assumed as a "measure" of the reciprocal *affinity* between the interacting proteins or, equivalently, as their *elective propensity* to realize a *stable* compound.

III. AN OSTENSIVE EXAMPLE

The example deals with diabetic therapy. It is well-known that human insulin has a reduced affinity with blood albumin, so that the subcutaneously injected insulin cannot efficiently be conveyed by blood albumin in the various parts of the body.

The therapy then consists in adopting a modified form of insulin, which presents a higher affinity with blood albumin. The modified form of insulin usually adopted is insulin detemir, also termed as levemir.

Figure 1 represents the three-dimensional structure of human insulin (51 amino acids), obtained by means of an appropriate simulator, run on a simple PC (10^9 Flops), in less than 1 s.

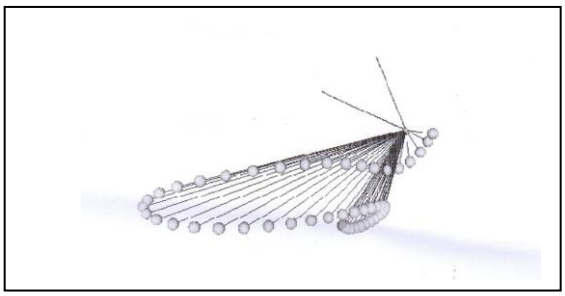

Figure 1 - Three-dimensional structure of human insulin (51 amino acids: 21 in subunit A and 30 in subunit B)

The simulator was termed as Emerging Quality Simulator (EQS), precisely because based on the MOP and its corresponding "emerging solutions" (see later on). The 3D structure so obtained can easily be modified (if needed) by means of slightly variations of some parameters of the RS.

This allows us to achieve a more accurate comparison, not only with the spatial configurations available in Literature (e.g., at the level of secondary structure), but also, and especially, with X-Ray Crystallography and/or Nuclear Magnetic Resonance (NMR) images available in qualified Protein Data Banks. This is because the output of the simulator, apart from the 3D structure, also gives the associated coordinates of all the amino acids, together with some other important indicators. Among others, and in particular, the corresponding "virtual work" associated to the protein.

Figure 2, in turn, represents the three-dimensional structure of blood albumin, made up of 585 amino acids. This spatial configuration was also obtained by means of the same simulator, run on the same PC, in a computation time of about 1 s.

As in the previous case, such a 3D structure can easily be compared with the corresponding spatial configurations available both in Literature and in Protein Data Banks.

At this stage, if we consider the interaction process between the two afore-mentioned proteins, we obtain that: i) insulin and albumin result as being characterized by virtual works whose values, expressed in the scale units usually adopted in EQS, are W_1 = 88.38 and W_2 = 587.66, respectively; ii) whereas the virtual work associated to the resulting compound is W_3 = 683.65.

Consequently, the corresponding ratio (3) gives

$$\delta W / W_3 = 0.0112 \qquad (4)$$

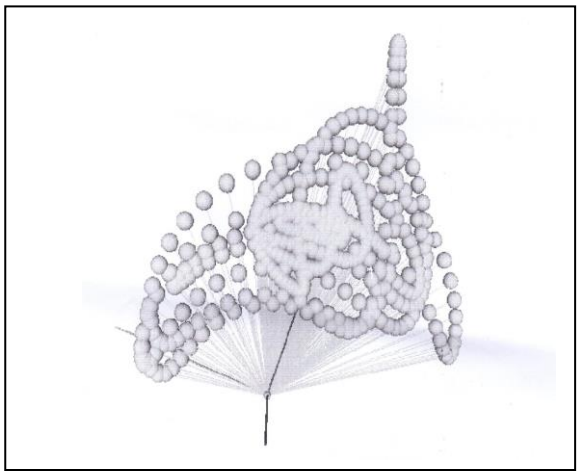

Figure 2 - Three-dimensional structure of blood albumin (585 amino acids)

This result clearly shows that human insulin has a very reduced affinity with albumin (about 1%). At the same time, it also explains why human albumin is usually modified in the form of levemir, in order to achieve a higher affinity.

Levemir insulin differs from human insulin in that the amino acid in position B30 is omitted, and a C14 fatty acid chain (termed as myristic acid) is attached to the amino acid B29.

Figure 3 represents the 3D structure of levemir, whose virtual work now results as being W_1^* = -29.95.

The *negative* value obtained simply indicates that the modified protein has an *inverse chirality* with respect to its primary form of insulin. This aspect generally favors the interaction process. In fact, the virtual work associated to the resulting compound now becomes W_3^* = 667.29.

Consequently, the interaction process between levemir and albumin (obtained by means of the same simulator in less than 2 s) gives origin to a final compound characterized by a higher "excess" of virtual work (2).

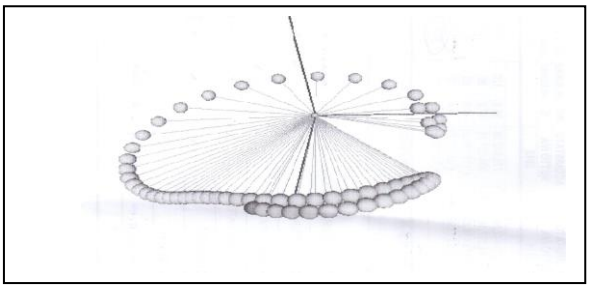

Figure 3 - Three-dimensional structure of levemir (50 amino acids plus the chain of 14 atoms of C)

Correspondingly, ratio (3) now gives

$$\{W_3^* - (W_1^* + W_2)\}/(W_1^* + W_2) = 0.1965 \quad (5)$$

This result clearly shows that such a modified form of human insulin presents an affinity of about 20% with respect to blood albumin. A value that allows levemir to be conveyed by albumin, without preventing, however, its subsequent release in the various parts of the body.

IV. INFORMATICS ADVANCES

The improvements here considered are directly referable to the *formal properties* that are intrinsic to the mathematical models adopted. In fact, any system modeled on the basis of the MOP, always presents *explicit solutions* in terms of *Incipient Differential Calculus* (see [3] and [4]).

This means that the method proposed has the *capacity* of *predicting the 3D structure of the resulting compound* essentially because the latter is understood as a self-organizing system of *ordinal nature*, and thus as *intrinsically* "*irreducible*" to functional relationships between its parts.

This correlatively also means: i) *a reduced number of computations*; ii) *a reduced need* of *High Performance Computing* (HPC); iii) *a reduced incidence* of *special numerical methods* to be adopted to get the corresponding solution.

What's more, the explicit solutions so obtained can also be termed as "*emerging solutions*" (see [5] and [8]), because *they always show an ordinal information content which is much higher than the corresponding content of the initial formulation of the problem.*

This is because the MOP is specifically finalized to describe "self-organizing" systems according to a *holistic approach*, in which, as is well-known, "*the whole is much more than the sum of its parts*".

V. BIO-INFORMATICS IN THE LIGHT OF THE MAXIMUM ORDINALITY PRINCIPLE

The method of solution previously illustrated with specific reference to Protein-Protein Interaction is also applicable to the majority of biological problems usually dealt with through informatics methods. In this sense, PPI only represents an ostensive example.

The same approach, in fact, has previously been adopted to improve the efficiency of the *exon skipping* method, usually used in Duchenne Muscular Dystrophy [9].

In such a case, the method enabled us to select the most appropriate Antisense Oligo-Nucleotides (AONs) with reference to four specific Exons: 51, 48, 44 and 39.

The pertinent experimental tests (in vitro and in vivo) are still in progress at LUMC (Leiden University Medical

Center) and the corresponding final results will be available at the end of next May.

This would indicate that the methodology here proposed could also be adopted in the case of Molecular Docking and Drug Design. In fact, it allows us to choose the *optimal ligand*, that is the one which is characterized by the most appropriate researched affinity (3), as in the case of exon skipping in DMD, previously mentioned.

Consequently, when considered from a more general point of view, the paper would intend to show that, in the light of the MOP, it is possible to realize mathematical models of several biological Systems, with very significant related advantages.

VI. CONCLUSION

The methodology here proposed seems to be able to give a significant contribution to Pharmacology. This is because the results previously shown indicate that the dynamic evolutions of a wide variety of biological processes can adequately be described by adopting the same reference principle (namely, the MOP).

Consequently, the various biological processes to be analyzed, when modeled by means an appropriate simulator, can be run on a simple personal computer and, in addition, in a computation time of a few seconds (or one minute, at the most).

This means that, by adopting the afore-mentioned approach, any researcher would be able to analyze the dynamic behavior of any biological process of interest by means of his/her own PC, simply sitting at his/her own desk.

The solutions obtained, in fact, will always describe a System whose parts are related to each other according to *ordinal relationships*. In other words, according to the same "relationships" that precisely take origin from *generative processes*, such as, for instance, the *genesis of two brothers*.

"Brothers", in fact, are properly defined as such, not because of their *direct* relationships. That is: because they respect each other or they love each (in fact, they might also hate each other). They are "brothers", *in essence*, because generated by the same father (or the same mother, or both). That is, because of their *direct relationship* with the *generative cause* of their being born.

Such a *genetic* relationship represents in fact something that is *unique*, *specific* and *irreducible*. Consequently, they cannot simply be accounted for as "two" (1+1), but as *one sole entity* (that is, as a whole), in spite of their clear reciprocal distinction. Consequently, the proper meaning of "*brothers*" refers to a clear "*irreducible extra*".

Precisely that represented by *their specific relationship* with the *same genetic* principle.

In accordance with such a concept of *ordinal relationship*, in the case of a given protein the *direct* relationship between two *any* amino acids is considered as being of the second order.

The *first order* relationship, in fact, is that which relates all the amino acids to the *same generative activity* of the protein, always understood as a *whole*.

This is why the explicit solutions that "emerge" from the MOP immediately give the positions of all the amino acids with respect to the internal axes of the protein (see Introduction, in particular, points i) and ii)).

The same concept is evidently valid for any *self-organizing* system, when described on the basis of the MOP.

The formal enunciation of this principle, in fact, first given in [5], is nothing but the reformulation of the Maximum Em-Power Principle, proposed by H.T. Odum in [10][11][12], understood as an updated version of the Fourth Thermodynamic Principle, first enunciated by Boltzmann and afterwards by Lotka, in [13] and [14], respectively.

Odum's enunciation, in fact, after having received an appropriate mathematical formulation *under dynamic conditions* in [15], was reformulated in more general *terms* in [5], by means of *a new concept of derivative*, the "*incipient*" *derivative*, whose mathematical definition was first introduced in [3] and further developed in [4].

The corresponding verbal enunciation of the MOP then became: "*Every system tends to maximize its own ordinality, including that of the surrounding habitat*".

REFERENCES

[1] H. Poincaré, "Les Méthodes Nouvelles de la Mécanique Céleste", 1889. Ed. Librerie Scientifique et Technique A. Blachard. Vol. I, II, III, Paris, 1987.

[2] C. Giannantoni, "From Transformity to Ordinality, or better: from Generative Transformity to Ordinal Generativity". Proceedings of the 5th Emergy Conference. Gainesville, Florida, USA, January 31-February 2, 2008, pp. 581-598.

[3] C. Giannantoni, "The Problem of the Initial Conditions and Their Physical Meaning in Linear Differential Equations of Fractional Order". Applied Mathematics and Computation 141, 2003, pp. 87-102.

[4] C. Giannantoni, "Mathematics for Generative Processes: Living and Non-Living Systems". Applied Mathematics and Computation 189, 2006, pp. 324-340.

[5] C. Giannantoni, "The Maximum Ordinality Principle. A Harmonious Dissonance". Proceedings of the 6th Emergy Conference. Gainesville, USA, January 14-16, 2010, pp. 55-72.

[6] C. Giannantoni, "Protein Folding, Molecular Docking, Drug Design. The Role of the Derivative "Drift" in Complex Systems Dynamics". Proceedings of the Third International Conference on Bioinformatics. Valencia, Spain, January 20-24, 2010, pp. 193-199.

[7] C. Giannantoni, “Bio-Informatics in the Light of the Maximum Ordinality Principle. The Case of Duchenne Muscular Dystrophy”. Proceedings of the 4th International Conference on Bioinformatics. Rome, January 26-29, 2011, pp. 244-250.

[8] C. Giannantoni, “The Relevance of Emerging Solutions for Thinking, Decision Making and Acting. The case of Smart Grids”. Ecological Modelling 271, 2014, pp. 62-71.

[9] A. Aartsma-Rus, W. E. Kaman, R. Weij, J. T. Den Dunnen, G. J. van Ommen, J. C. van Deutekom, “Exploring the Frontiers of Therapeutic Exon Skipping for Duchenne Muscular Dystrophy by Double Targeting within One or Multiple Exons”. Molecular Therapy, Vol. 14, 2006, pp. 401-407.

[10] H. T. Odum, “Ecological and General Systems. An Introduction to Systems Ecology”. Re. Edition. University Press Colorado, 1994.

[11] H. T. Odum, “Environmental Accounting”. Environmental Engineering Sciences. University of Florida, 1994.

[12] H. T. Odum, “Self Organization and Maximum Power”. Environ. Engineering Sciences. University of Florida, 1994.

[13] L. Boltzmann, “Der zweite Haupsatz der mechanishen warme Theorie“. Almanach der K. Acad. Wiss Mechanishe, Wien, 1905, Vol. 36, pp. 225-299 (printing of a lecture given by Boltzman in 1886).

[14] A. J. Lotka, “The Law of Evolution as a Maximal Principle”. Human Biology, a record of research. September 1945, Vol. 17, n. 3.

[15] C. Giannantoni, “The Maximum Em-Power Principle as the basis for Thermodynamics of Quality”. Ed. S.G.E., Padua, 2002, ISBN 88-86281-76-5.